



Addor, N., Do, H., Coxon, G., Alvarez-Garretón, C., Fowler, K., & Mendoza, P. (2019). Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges. *Hydrological Sciences Journal*. <https://doi.org/10.1080/02626667.2019.1683182>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1080/02626667.2019.1683182](https://doi.org/10.1080/02626667.2019.1683182)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Taylor and Francis at <https://doi.org/10.1080/02626667.2019.1683182> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges

Nans Addor, Hong X. Do, Camila Alvarez-Garreton, Gemma Coxon, Keirnan Fowler & Pablo A. Mendoza

To cite this article: Nans Addor, Hong X. Do, Camila Alvarez-Garreton, Gemma Coxon, Keirnan Fowler & Pablo A. Mendoza (2019): Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges, Hydrological Sciences Journal, DOI: [10.1080/02626667.2019.1683182](https://doi.org/10.1080/02626667.2019.1683182)

To link to this article: <https://doi.org/10.1080/02626667.2019.1683182>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 03 Dec 2019.



Submit your article to this journal [↗](#)



Article views: 449



View related articles [↗](#)



View Crossmark data [↗](#)

Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges

Nans Addor^a, Hong X. Do^{b,c,d}, Camila Alvarez-Garretón^{e,f}, Gemma Coxon^g, Keirnan Fowler^h and Pablo A. Mendoza^{i,j}

^aClimatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, UK; ^bSchool of Civil, Environmental and Mining Engineering, University of Adelaide, Adelaide, Australia; ^cFaculty of Environment and Natural Resources, Nong Lam University, Ho Chi Minh City, Vietnam; ^dSchool for Environment and Sustainability, University of Michigan, Ann Arbor, MI, USA; ^eInstituto de Conservación, Biodiversidad y Territorio, Universidad Austral de Chile, Valdivia, Chile; ^fCenter for Climate and Resilience Research, Santiago, Chile; ^gSchool of Geographical Sciences, University of Bristol, Bristol, UK; ^hDepartment of Infrastructure Engineering, University of Melbourne, Parkville, Australia; ⁱDepartment of Civil Engineering, Universidad de Chile, Santiago, Chile; ^jAdvanced Mining Technology Center, Universidad de Chile, Santiago, Chile

ABSTRACT

Large-sample hydrology (LSH) relies on data from large sets (tens to thousands) of catchments to go beyond individual case studies and derive robust conclusions on hydrological processes and models. Numerous LSH datasets have recently been released, covering a wide range of regions and relying on increasingly diverse data sources to characterize catchment behaviour. These datasets offer novel opportunities, yet they are also limited by their lack of comparability, uncertainty estimates and characterization of human impacts. This article (i) underscores the key role of LSH datasets in hydrological studies, (ii) provides a review of currently available LSH datasets, (iii) highlights current limitations of LSH datasets and (iv) proposes guidelines and coordinated actions to overcome these limitations. These guidelines and actions aim to standardize and automatize the creation of LSH datasets worldwide, and to enhance the reproducibility and comparability of hydrological studies.

ARTICLE HISTORY

Received 31 March 2019
Accepted 26 September 2019

EDITOR

A. Castellarin

GUEST EDITOR

C. Cudennec

KEYWORDS

streamflow records; data standardization; reproducibility of hydrological experiments; data uncertainties; human interventions; cloud computing

1 Introduction: from comparative hydrology to large-sample hydrology

Large-sample hydrology (LSH) makes use of datasets involving large sets of catchments to derive robust conclusions on hydrological processes and models. LSH finds its roots in the field of comparative hydrology (Kovács 1984, Falkenmark and Chapman 1989), whose foundations were set in the framework of the International Hydrological Programme, launched by UNESCO in 1975. The general motivation of comparative hydrology is to learn from hydrological similarities and differences between places around the world, and interpret these in terms of underlying climate-landscape-human controls (e.g. McMahon 1982, Finlayson *et al.* 1986, Peel *et al.* 2001, 2004, Sivapalan 2009, Troch *et al.* 2009, Thompson *et al.* 2011). At that time, a key objective was to facilitate transfer of knowledge between regions, and for instance, to determine to what extent available hydrological theories and models, which were derived mostly for temperate regions of Europe and North America, could be applied in other regions (Falkenmark and Chapman 1989).

LSH follows similar objectives but puts a stronger emphasis on the need to establish robust principles by leveraging large sets of catchments, which led to the name “large-sample hydrology”. In this paper, we use the word “sample” more often than “set” as, in our view, the former better conveys the

idea that the basins we work with are drawn from a wide range of hydrological conditions and should enable us to formulate conclusions about basins we have not sampled. Andréassian *et al.* (2006a) underscore that model intercomparisons should be based on a significant number of catchments to deliver robust conclusions that are not the result of chance. Similarly, Gupta *et al.* (2014) insist that general hydrological principles should be derived from statistically significant relationships, which are unobtainable with data from only a few catchments. This makes LSH a branch, rather than a replacement, of comparative hydrology, and thus several comparative hydrology investigations can also be classified as LSH research (e.g. Singh *et al.* 2014).

Alongside large-sample hydrology, large-scale hydrology has become established (e.g. Cloke and Hannah 2011, Wood *et al.* 2011, Bierkens 2015). These two fields are complementary, as they both provide generalizable knowledge on the terrestrial water cycle across a range of hydroclimatic conditions. A notable difference between them lies in the scale and spatial continuity of the area covered. A large sample of catchments can cover a vast area, but this area is made of separate catchments. In contrast, large-scale hydrology explores “*spatial scales greater than a single river basin all the way up to the entire planet*” to use the definition of Cloke and Hannah (2011). Further, while streamflow measurements are

a cornerstone of catchment hydrology and LSH, at larger spatial scales, the focus is traditionally on other fluxes (e.g. evapotranspiration) and state variables (e.g. soil moisture). The gap between these two fields is however quickly reducing, with the development of gridded streamflow observations (Fekete *et al.* 2002, Gudmundsson and Seneviratne 2016, Ghiggi *et al.* 2019), ever larger domains covered by rainfall-runoff models (Beck *et al.* 2016), the ever finer resolution of large-scale models (Wood *et al.* 2011, Bierkens *et al.* 2015), assessments of the influence of catchment-scale processes on the performances of large-scale models (Kauffeldt *et al.* 2016, Fang *et al.* 2017, Veldkamp *et al.* 2018, Zaherpour *et al.* 2018), and the inclusion of streamflow simulations from macroscale models in LSH investigations (Rakovec *et al.* 2016, Zink *et al.* 2017, Do *et al.* 2019).

In this paper, we focus on LSH and, more specifically, on datasets providing streamflow data for a large number of catchments. Such datasets form the foundation of a wide range of hydrological studies dedicated to catchment classification (e.g. Sawicz *et al.* 2011, Kuentz *et al.* 2017, Knoben *et al.* 2018), extreme events (e.g. Mallakpour and Villarini 2015, Tjiedeman *et al.* 2016, Berghuijs *et al.* 2017, Blöschl *et al.* 2017, Do *et al.* 2017, Gudmundsson *et al.* 2019), terrestrial water storage (e.g. Zhang *et al.* 2017), data and model uncertainties (e.g. McMillan *et al.* 2012, Coxon *et al.* 2015, Beck *et al.* 2017), hydrological model evaluation and benchmarking (e.g. Mathevet *et al.* 2006, Andréassian *et al.* 2009, Gudmundsson *et al.* 2012, Coron *et al.* 2012, Coxon *et al.* 2013, Fowler *et al.* 2016, McMillan *et al.* 2016a, Newman *et al.* 2017, Seibert *et al.* 2018, Kratzert *et al.* 2019), parameter estimation of hydrological models (e.g. Perrin *et al.* 2008, Oudin *et al.* 2010, Andréassian *et al.* 2014, Beck *et al.* 2016, Rakovec *et al.* 2016, Hirpa *et al.* 2018), regionalization using machine learning algorithms (Beck *et al.* 2015, Addor *et al.* 2018, Barbarossa *et al.* 2018, Kratzert *et al.*, 2018, Kratzert *et al.*, 2019), human impacts on hydrology (e.g. Alvarez-Garreton *et al.* 2018, Tjiedeman *et al.* 2018a, 2018b), streamflow forecasting (e.g. Harrigan *et al.* 2018, Slater and Villarini 2018) and climate change impacts assessments (e.g. Melsen *et al.* 2018). LSH datasets underpin key advances in hydrological sciences and are fundamental to major community-wide efforts, in particular to the Prediction in Ungauged Basins (PUB, Hrachowitz *et al.* 2013) and Panta Rhei (Montanari *et al.* 2013, McMillan *et al.* 2016b) initiatives of the International Association of Hydrological Sciences (IAHS).

The diversity and content of LSH datasets is expanding rapidly. Gupta *et al.* (2014) highlighted several datasets potentially useful for LSH applications and, since then, several datasets dedicated to LSH have been published. They cover a far greater number of catchments, hydroclimatic regions and catchment attributes than what was available just a few years ago. In Section 2, we provide a snapshot of this development and give an overview of LSH datasets currently available. These recent advances and the opportunities they offer are remarkable, yet, as creators and users of LSH datasets, we argue that it is now crucial to better coordinate the production and exchange of LSH datasets worldwide. For this *Hydrological Sciences Journal* special issue on “Hydrological data:

opportunities and barriers”, we identified four LSH challenges that require immediate attention: (i) the difficulties of inter-dataset comparison, (ii) the lack of uncertainty estimates, (iii) the insufficient representation of human interventions, and (iv) the still limited accessibility of hydrological observations. These challenges are discussed in Section 3. We then list simple, concrete actions (Section 4) and outline coordinated efforts (Section 5) to overcome these barriers. Conclusions are presented in Section 6.

2 Recent progress in the development of LSH datasets

In this section, we review LSH datasets currently available, and focus on LSH datasets fulfilling two criteria, referred to below as “minimum requirements”: (a) the dataset must contain streamflow observations and (b) basic identifiers for each stream-gauge (i.e. name, catchment area, gauge coordinates) must be included. We did not set a specific number of catchments to define a sample as “large”, as the needs of each study are unique. For instance, tens of carefully selected catchments can enable insightful regional comparisons (e.g. Bennett *et al.* 2018, Burn and Whitfield 2018, Fowler *et al.* 2018), while one may argue that thousands of catchments are needed for global scale investigations (e.g. van Dijk *et al.* 2013, Beck *et al.* 2015, Do *et al.* 2017, Gudmundsson *et al.* 2019). In addition, this paper focuses on datasets available in digital form with relative ease of access. It does not cover individual national water archives, the classical data source resulting from national-scale streamflow monitoring, as some of them are only maintained in paper form or subject to strict data-distribution policies. However, these national archives form the basis of the LSH datasets described below.

2.1 Data available through LSH datasets

The nature of the data covered by LSH datasets varies significantly from one dataset to the next. To facilitate the navigation and selection of LSH datasets by potential users, we classify these data into three categories: (i) streamflow observations, (ii) hydrometeorological time series and (iii) landscape and hydroclimatic attributes (Table 1).

Streamflow observations is a category on its own, since we make their availability a minimum requirement for the dataset to be considered here. Some LSH datasets complement streamflow observations with other hydrometeorological time series, such as precipitation and temperature. Further, variables characterizing the landscape of the catchments, for instance their land cover or soil, are included in some datasets. We note that the availability of hydrometeorological time series and catchment landscape attributes varies strongly among LSH datasets. The wealth of available spatial data (e.g. gridded meteorological observations or remotely sensed vegetation products) means that LSH datasets creators only select and process a subset of available datasets. As a result, different LSH datasets are best adapted to different research pursuits. For example, LSH datasets including atmospheric forcing time series for each catchment (Schaafe *et al.* 2006, Newman *et al.* 2015a, Addor *et al.* 2017, Arsenault *et al.* 2016, Alvarez-Garreton *et al.* 2018) are well suited for hydrological modelling, whereas those

Table 1. Summary of the data potentially included in LSH datasets.

	Minimum requirements defined in this paper	Additional information	Metadata and uncertainty estimates
Streamflow observations and basic catchment identifiers	Daily or monthly streamflow records Basic identifiers (name, gauge coordinates, catchment area)	Hydrological signatures GIS layer of catchment boundary	Quality flags for observations Flag if station included in other LSH or 'reference' (high quality) datasets References to nested or neighbour catchments Rating curve information Infos on gauge relocation Uncertainty estimates
Hydrometeorological time series	No minimum requirements	Precipitation Temperature Evapotranspiration Soil moisture Snow water equivalent	Discussion of data sources limitations Uncertainty estimates
Landscape and hydroclimatic attributes	No minimum requirements	Topography Climate Land cover Soil Geology Stream network Human impacts (abstractions, irrigation, dams, human population)	Discussion of data sources limitations Uncertainty estimates

including catchment attributes representing human presence or irrigation (Do *et al.* 2018a, Gudmundsson *et al.* 2018a) are better adapted for detection and attribution studies.

In addition, some datasets provide metadata and uncertainty estimates. For example, catchment boundaries may be provided with quality flags, and time series may be subject to a homogeneity assessment to produce uncertainty estimates (Do *et al.* 2018a, Gudmundsson *et al.* 2018a). Some datasets derived meteorological time series using several data-products to reflect forcing uncertainty (Newman *et al.* 2015a, Alvarez-Garreton *et al.* 2018).

2.2 LSH datasets currently available

Table 2 provides an overview of eleven key LSH datasets. These datasets cover different parts of the world and include basins from a single country to the entire globe. The access to these datasets is unrestricted for scientific purposes. However, the licensing policies vary, with some datasets being fully available in the public domain, while others requiring data requests in written form.

At the global scale, the Global Runoff Data Base (GRDB) is arguably the main dataset used for streamflow investigations, including LSH studies. This database is maintained by the Global Runoff Data Centre (GRDC), which operates under the auspices of the World Meteorological Organization (WMO) since 1988 and holds records of daily and monthly streamflow across more than 9000 stations globally (GRDC, 2015). This global initiative is supported voluntarily by national authorities and thus, data contributions depend on

the capacity of corresponding agencies. As a result, some countries are sparsely represented in GRDB, even though data of reasonable quality are available (e.g. most stations in Asia have not been updated since the 1990s; GRDC 2015). To facilitate access to streamflow data from stations across the world, the Global Streamflow Indices and Metadata (GSIM) archive was recently produced (Do *et al.* 2018a, Gudmundsson *et al.* 2018a). GSIM is an expansion of GRDB, which was produced by collating streamflow observations from 11 other publicly available databases (including three LSH datasets also described in Table 2) and publishing standardized metadata relevant to LSH research (Do *et al.* 2018b). To make hydrological information publicly available, even when raw data cannot be redistributed, GSIM contains time series of streamflow indices at different temporal resolutions (i.e. monthly/seasonal/yearly timestep) derived from raw daily records (Gudmundsson *et al.* 2018b).

At the continental scale, the European Water Archive (EWA) is one of the most comprehensive streamflow time-series archives with records of more than 3000 river gauging stations contributed by 29 European national hydrological services. The EWA is now hosted by GRDC and can be accessed under the GRDC data policy. However, EWA has not been updated since 2014 and no future updates are planned by GRDC. Since then, some national hydrological services have allowed GRDC to integrate EWA stations into GRDB, so records for these stations are now regularly updated through new releases of GRDB. EWA streamflow records were recently combined with GRDB stations and the European catchments from the Hydrological Predictions for the Environment (E-HYPE) model and

Table 2. Key LSH datasets currently available. Q: streamflow; P: precipitation; T: temperature; PET: potential evapotranspiration; SWE: snow water equivalent.

Name	No. of catchments	Location	Hydrometeorological time series	Basic catchment identifiers	Catchment attributes	Availability	DOI or URL	Reference or provider
Global Streamflow Indices and Metadata Archive (GSIM)	35 002	Global	Monthly, seasonal and annual streamflow indices	Name, gauge coordinates, catchment area and boundary	Topography, land cover, geology, irrigation, human population, soil	Freely available under Creative Commons Attribution-ShareAlike 3.0 Unported license	https://doi.org/10.1594/PANGAEA.887477	Do <i>et al.</i> (2018a); Gudmundsson <i>et al.</i> (2018a)
Global Runoff Data Base (GRDB)	9543	Global	Q	Name, gauge coordinates, catchment boundary available for some stations	No catchment attributes available	Free and unrestricted (but identified) access. No redistribution of the data by the user.	https://www.bafg.de/GRDC/	GRDC
European Water Archive (EWA)	3731	Europe	Q	Name, gauge coordinates, catchment boundary available for some stations	No catchment attributes available	Free and unrestricted (but identified) access. No redistribution of the data by the user.	http://www.bafg.de/GRDC/EN/04_spcldtbs/42_EWA/ewa_node.html	GRDC
European catchments of Hydrological Predictions for the Environment model (E-HYPE)	35 215 catchments 1366 river gauges	Europe	Q	Name, gauge coordinates, catchment area and boundary	Topography, climatic indices, hydrological signatures, land cover, geology, soil	Freely available under Creative Commons Attribution Share Alike 4.0 International license	https://doi.org/10.5281/zenodo.581428	Kuentz <i>et al.</i> (2017)
							https://doi.org/10.5281/zenodo.581429	
							https://doi.org/10.5281/zenodo.581430	
							https://doi.org/10.5281/zenodo.581431	
							https://doi.org/10.5281/zenodo.581432	
							https://doi.org/10.5281/zenodo.581434	
							https://doi.org/10.5281/zenodo.581435	
MModel Parameter Estimation Experiment (MOPEX)	438	US	P, T, Q	Name, gauge coordinates, catchment area and boundary	Topography, climatic indices, land cover, soil	Freely available	http://www.nws.noaa.gov/oh/mopex/	Schaake <i>et al.</i> (2006)
CANadian mOdel Parameter EXperiment (CANOPEX)	698	Canada	P, T, Q	Name, gauge coordinates, catchment area and boundary	No catchment attribute available	Freely available for non-commercial applications	http://canopex.etsmtl.net	Arsenault <i>et al.</i> (2016)
Catchment Attributes and Meteorology for Large-sample Studies (CAMELS)	671	US	P, T, Q, PET	Name, gauge coordinates, catchment area and boundary	Topography, climatic indices, hydrological signatures, land cover, soil, geology	Freely available under Creative Commons Attribution 4.0 International	https://doi.org/10.5065/D6MM2F4D	Newman <i>et al.</i> (2015a); Addor <i>et al.</i> (2017)
Catchment Attributes and Meteorology for Large-sample Studies – Chile (CAMELS-CL)	516	Chile	P, T, Q, PET, SWE	Name, gauge coordinates, catchment area and boundary	Topography, climatic indices, hydrological signatures, land cover, geology, water use	Freely available under Creative Commons Attribution 4.0 International	https://doi.org/10.1594/PANGAEA.894885	Alvarez-Garreton <i>et al.</i> (2018)

(Continued)

Table 2. (Continued).

Name	No. of catchments	Location	Hydrometeorological time series	Basic catchment identifiers	Catchment attributes	Availability	DOI or URL	Reference or provider
Regional hydrometeorological data network for the pan-Arctic Region (ARCTICNET)	139	Russia (streamflow data from other pan-Arctic regions available as monthly means) China	Q	Name, gauge coordinates, catchment area	No catchment attributes available	Freely available	http://www.r-arcticnet.sr.unh.edu/v4.0/index.html	The State Hydrological Institute, Russia
China Hydrology Data Project (CHDP)	163	China	Q	Name, gauge coordinates, catchment area	No catchment attributes available	Available upon request to Amanda Henck Schmidt	http://www2.oberlin.edu/faculty/aschmidt/chdp/index.html	Henck <i>et al.</i> (2010); Schmidt <i>et al.</i> (2011)
GEWEX Asian Monsoon Experiment – Tropics project (GAME)	129	Thailand	Q	Name, gauge coordinates, catchment area	No catchment attributes available	Freely available	http://hydro.iis.u-tokyo.ac.jp/GAME-T/GAIN-T/outline/rid-river/index.html	University of Tokyo, Japan

used to improve the characterization and understanding of hydrologic variability across Europe through catchment classification (Kuentz *et al.* 2017).

Another long-standing LSH dataset is the data archive of the Model Parameter Estimation Experiment project (MOPEX; Duan *et al.* 2006), which includes data for 438 catchments across the USA. In addition to hydrometeorological observations, MOPEX provides attributes for catchments representing different hydroclimatic conditions and was one of the main data sources underpinning the PUB decade (Andréassian *et al.* 2006b). However, MOPEX hydrometeorological time series stop in 2003 and MOPEX is no longer updated.

At the national scale, several datasets have been developed with an approach similar to MOPEX. The Catchment Attributes and MEteorology for Large-sample Studies dataset (CAMELS; Newman *et al.* 2015a, Addor *et al.* 2017) uses recent datasets to provide up-to-date hydrometeorological variables and a variety of landscape attributes for 671 catchments across the contiguous United States. CAMELS also includes detailed descriptions of the methods used to derive catchment attributes and a discussion of several of data-source caveats. A similar approach was used to produce the CAMELS-Chile dataset (CAMELS-CL; Alvarez-Garreton *et al.* 2018), which provides an overview of regional variations in hydroclimatic conditions over Chile and an assessment of human interventions to streamflow regime across 516 catchments. Meteorological and hydrologic data for 698 catchments in Canada are available through the CANOPEX database (Arsenault *et al.* 2016).

3 Limitations of current LSH datasets

To guide the development of future LSH datasets, in this section, we highlight four typical limitations of LSH datasets released so far (i) the lack of common standards impedes the comparison of basins from different datasets, (ii) the lack of uncertainty estimates prevents users from assessing data reliability, (iii) the extent of human interventions is rarely characterized, and (iv) data accessibility is still limited.

3.1 The lack of common standards impedes the comparison of basins from different datasets

Comparative hydrology is only possible if the data from different catchments are consistently processed, and thus can be compared. While the comparison of catchments from the same LSH dataset is usually straightforward, comparisons across LSH datasets is often challenging because different naming conventions, data sources and methods for calculating the same variables are used from one dataset to the next. This issue is part of the wider challenge of using common standards and protocols when producing and processing environmental data (e.g. Horsburgh *et al.* 2009, Ceola *et al.* 2015), and it critically limits our ability to combine and learn from several LSH datasets.

3.2 The lack of metadata and uncertainty estimates prevents users from assessing data reliability

When using data from many catchments, assessing data errors is key, as they can bias comparisons between catchments. Yet,

there is still a clear lack of uncertainty estimates accompanying LSH datasets. Uncertainties in atmospheric forcings are receiving the most attention and are increasingly characterized by relying on several datasets (e.g. Newman *et al.* 2015a, Alvarez-Garreton *et al.* 2018). In contrast, uncertainties in catchment attributes (e.g. land cover, soil characteristics) are rarely quantified, or even acknowledged, in LSH datasets.

Streamflow uncertainty estimates and metadata on gauge information are also rarely available, although the limitations and uncertainties of streamflow time series are well known (e.g. McMillan *et al.* 2012). Streamflow metadata are often not available due to management practices of data providers (Hannah *et al.* 2011), the loss of metadata during data transfers from providers to international data archives or poor upkeep of this information (Gudmundsson *et al.* 2018a). Further, even when metadata are available, assessing streamflow uncertainties across large samples of catchments remains a challenge, as different methods are recommended for different gauge types (Kiang *et al.* 2018).

3.3 The extent of human interventions is rarely characterized

LSH datasets have historically focused on physical attributes, making use of the wealth of data currently available to characterize hydrological behaviour (Tables 1 and 2). In comparison, human interventions are still poorly characterized in LSH datasets, although human alterations have large impacts on the natural water cycle (e.g. Vörösmarty *et al.* 2000, Hanasaki *et al.* 2006). These impacts may be comparable to climate change effects at the regional scale (Ferguson and Maxwell 2012) and threaten sustainability at the global scale (Jaramillo and Destouni 2015). For example, increased reservoir storage not only affects runoff seasonality but also the frequency of low/high flow events observed at the catchment outlet (e.g. Wehren *et al.* 2010), and changes in land cover influence the distribution of streamflow, specifically baseflow volumes and flashiness of runoff (e.g. Vertessy 2000, Brown *et al.* 2005, Alvarez-Garreton *et al.* 2019). Consequently, providing information on such alterations is critical to assess the magnitude of human impacts on hydrological behaviour (e.g. Alvarez-Garreton *et al.* 2018) and to incorporate human interventions in hydrological models (e.g. Payan *et al.* 2008, Liu *et al.* 2017, Veldkamp *et al.* 2018), thus enhancing our ability to provide reliable hydrological simulations in an increasingly human impacted environment.

3.4 LSH datasets are rarely FAIR – findable, accessible, interoperable and reusable

To advance LSH, progress is needed to make LSH datasets more FAIR (Findable, Accessible, Interoperable and Reusable, see Wilkinson *et al.* 2016 and the Open Data Charter 2015). Currently, many digitized datasets are stored in local repositories or data portals are unknown to data users (not “Findable”). Data accessibility is still limited for many regions of the world (not “Accessible”) biasing LSH studies towards countries with greater accessibility. LSH datasets are hosted in different locations with a range of different upkeeping practices (not “Interoperable”). The license of many hydrometeorological records does not allow users to share data in their possession (not “Reusable”).

Disparities in the availability of streamflow records worldwide are a significant barrier to the development of global LSH datasets. Figure 1 shows the varying temporal coverage across the globe, with stations in North America and Europe generally having the longest records. Importantly, “white space” still dominates in many regions of the world, as also shown in other studies (e.g. Barbarossa *et al.* 2018). In some cases, this can be attributed to the lack of stations, in particular in extreme environments. However, in several regions data streamflow records do exist but they are not accessible because (i) data are not available in digitized form, (ii) digitized data are hosted in a local repository and data authorities do not have the resources to process data requests, (iii) data are not made available or are subject to payable fees, and (iv) the one-station-at-a-time downloading process (mouse and keyboard interactions required) hampers data retrieval for LSH studies.

4 Guidelines for the production of LSH datasets

To overcome the limitations outlined in Section 3, we propose six simple guidelines to support the creation of future LSH datasets (presented in this Section) and four coordinated actions (presented in Section 5). The limitations, guidelines and actions are summarized graphically in Fig. 2.

The six guidelines outlined here are simple to follow and will improve the value and usability of future datasets. We consider them as minimum requirements to be satisfied by new LSH datasets, and hence suggest that they are checked by both LSH dataset creators and by reviewers of papers introducing new LSH datasets.

4.1 Provide basic data for each basin

Streamflow observations remain the cornerstone of LSH, and thus new LSH datasets should make these records available. For streamflow records subject to strict redistribution data policy, releasing streamflow indices at different temporal resolutions is an

alternative (e.g. Do *et al.* 2018a, Gudmundsson *et al.* 2018a). The metadata should at least include the name, unique identifier (ID), river and geographical coordinates of each stream-gauge, as well as the catchment area and elevation information. Providing a shapefile of the catchment boundary associated with each stream-gauge (see e.g. Lehner 2012) should also be prioritized, so that users can derive additional attributes or time series from global or regional data products. Using the same digital elevation data source for all the basins is recommended, HydroSHEDS (<http://www.hydrosheds.org>) and Viewfinder (<http://viewfinderpanoramas.org>) being popular choices at the global scale.

4.2 Follow established standards when naming variables

The observance of common standards, including the use of a controlled vocabulary, is essential to ensure the consistency and comparability of environmental datasets (e.g. Horsburgh *et al.* 2009, Vitolo *et al.* 2015; Moine *et al.* 2014). Consistent variable names across LSH datasets should be used to make new datasets easier to utilize by the community and to facilitate inter-dataset comparisons. This is in addition to metadata, which describe the methods and data sources used to compute each variable.

The activities of the Open Geospatial Consortium (OGC), a not for profit organization working on the development of open standards for the global geospatial community, are key to progress on this front. The OGC set of standards most relevant to LSH is WATER ML-2 (<http://www.waterrml2.org>), as it is dedicated to hydro-meteorological observations and measurements. Similar conventions also exist for climate variables (e.g. Climate and Forecast Community Metadata Standard, <http://cfconventions.org>). These standards form the basis of variable naming, yet they only cover a fraction of all the variables relevant to LSH. Hence, our recommendation is to build on these standards, consider naming decisions made in other LSH datasets and improve them, with the goal to create a set of variable names that can be used across LSH datasets.

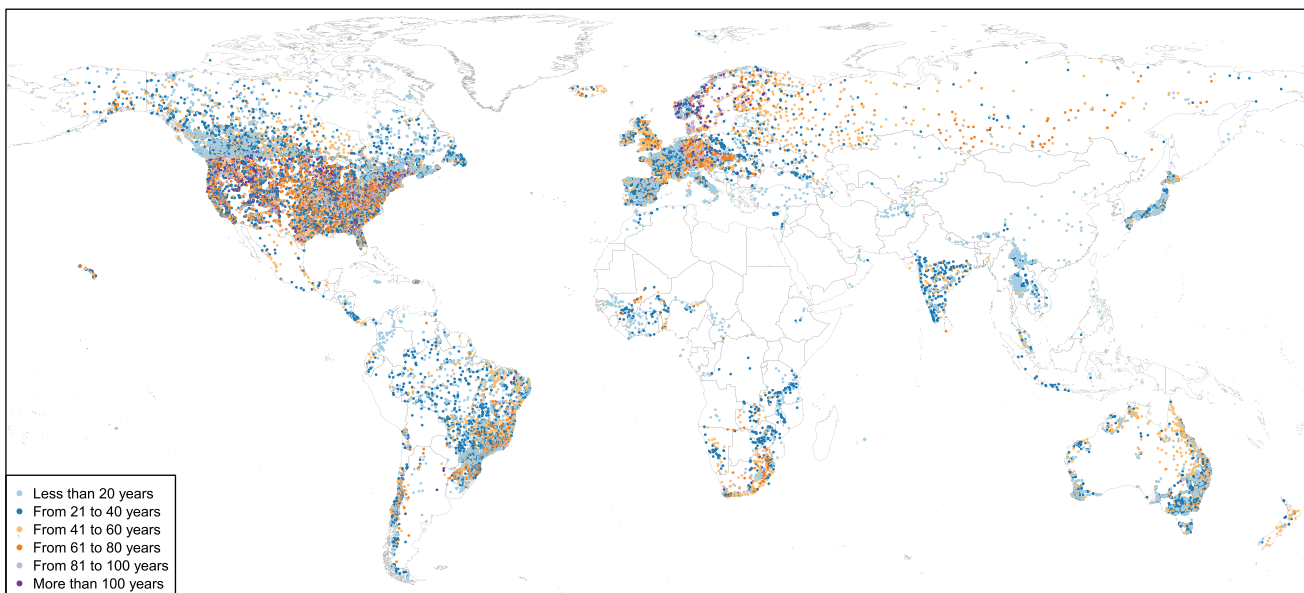


Figure 1. Overview of the location and record length of over 31,500 stream gauges from the Global Streamflow Indices and Metadata (GSIM) archive and the Catchment Attributes and MEteorology for Large-sample Studies dataset for Chile (CAMELS-CL, which was released after GSIM).

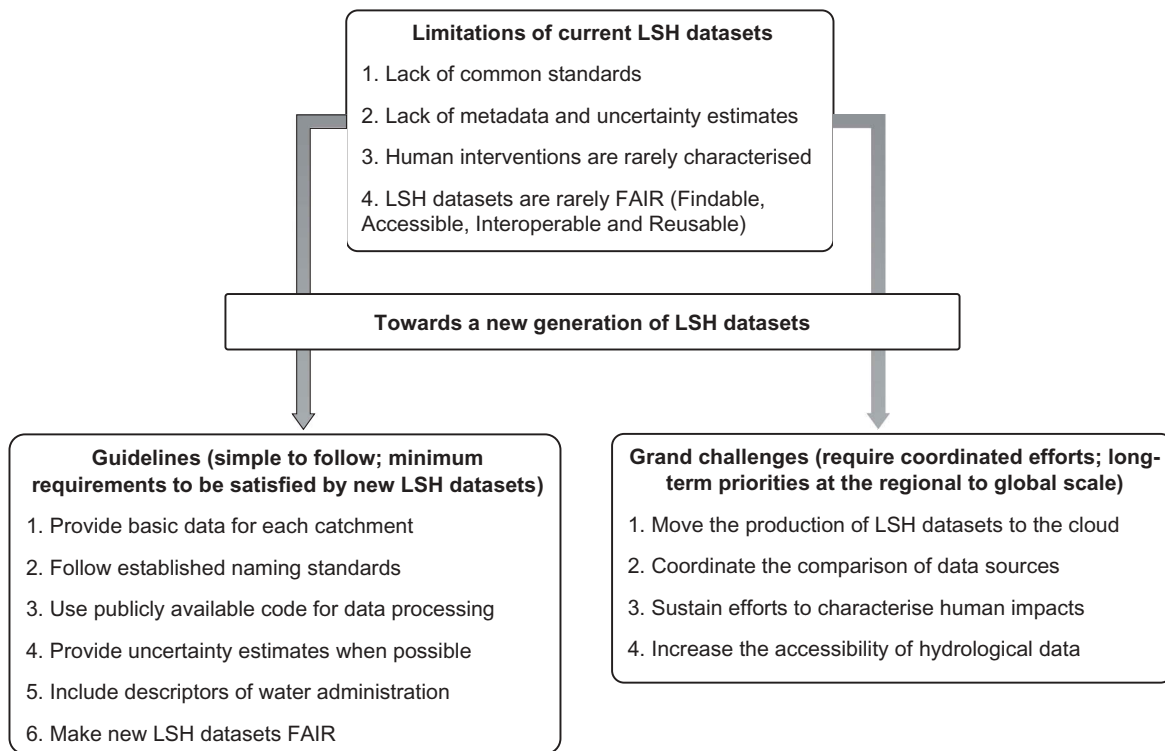


Figure 2. Schematic summary of the limitations of current LSH datasets (Section 3), proposed guidelines for new datasets (Section 4) and grand challenges for LSH (Section 5).

4.3 Use publicly available code for data processing

To improve transparency and reproducibility (Hutton *et al.* 2016), the code used for the creation of LSH datasets should be publicly available, either by using already-published code (e.g. packages, Greene and Thirumalai 2019) or making the code used available (e.g. on GitHub, Easterbrook 2014). Several packages and libraries already exist to compute key attributes (especially climate indices and hydrological signatures) in different programming languages (see <https://github.com/ropen-sci/hydrology> and Slater *et al.* 2019 for R). Given that hydrological signatures can be particularly sensitive to their formulation, as shown for instance by Stoezle *et al.* (2013) for recession coefficients, using publicly available code is essential.

Ideally, the shared code should cover more than the computation of climatic indices and hydrological signatures. It should, for instance, also include scripts to create catchment-averages from gridded products, and algorithms performing quality assurance tests of streamflow data. The goal is to create a library of scripts to perform standards LSH dataset creation tasks (such as those just mentioned), thereby increasing transparency and comparability. For instance, the scripts used to produce many attributes of the CAMELS dataset (Addor *et al.* 2017) are publicly available (<https://github.com/naddor/camels>) and have been used to produce the CAMELS-CL dataset (Alvarez-Garreton *et al.* 2018).

4.4 Provide uncertainty estimates for time series and catchment attributes

To allow users to assess the reliability of a LSH dataset, its quality should be evaluated and provided as metadata

alongside the dataset. Quality flags from data providers and simple numerical screening techniques can be used to develop quality assurance (QA) methods (see for example Gudmundsson *et al.* 2018a for flow QA procedures and Blenkinsop *et al.* 2017, Lewis *et al.* 2018 for rainfall QA procedures). This should be developed in cooperation with hydro-metric agencies who often employ QA procedures before the data is released. For streamflow observations, quality flags should be available for each gauge. Such information would help hydrologists to detect outlier basins and decide whether to include them in their analysis (e.g. Boldetti *et al.* 2010).

The uncertainty of data products used when producing a LSH dataset should be assessed. One opportunity for large sample hydrology is to construct multiple estimates of a given variable using different products or formulations. This is already evident in many of the LSH datasets highlighted in Section 2, Alvarez-Garreton *et al.* (2018) for example generated daily estimates of precipitation and potential evapotranspiration from multiple products, and is becoming more viable with the increasing availability of continental/global products (e.g. Beck *et al.* 2017).

4.5 Include descriptors of water administration systems

Water administration descriptors should be included in LSH datasets. Ideally, the following attributes should be provided at the most detailed spatial and temporal scales possible: (i) usage type (e.g. consumption, irrigation, hydropower, groundwater recharge, extraction), (ii) location, (iii) allocated volume, and (iv) timing. The first attribute indicates whether water returns to the rivers are expected, and hence should be completed by

additional information for attributes (ii)-(iv). Other information used as a proxy for human water use – such as catchment population, percentage of urban and agricultural land use, and the presence of dams (<http://globaldamwatch.org>) – are also valuable, particularly in regions where water usage data are not available.

4.6 Assess and increase the dataset FAIRness

For new LSH datasets to be findable, they should be documented in open-access, peer-reviewed journals and indexed via a DOI within publication databases. Within-agency technical reports are not sufficient to ensure findability. To be accessible, datasets should be downloadable from the internet at no cost and provide the option to download the entire dataset at once (in addition to site-by-site download). To be interoperable, the meaning of the data should be unambiguous regardless of the context. To be reusable, datasets need a licence allowing users to use, share, and build upon the existing dataset, encouraging collaboration and extension of existing datasets. We recommend to data owners to assess and increase the FAIRness of their dataset by using this online tool by the Australian National Data Service and partners (<https://www.andis-nectar-rds.org.au/fair-tool>, ANDS *et al.*, 2017).

5 Outlook: grand challenges and priorities for LSH

In this section, we discuss tasks that go beyond what can be expected from an individual LSH study, and require coordinated efforts from the LSH community, and in some cases, the wider geoscience community. We deliberately kept their formulation general and not prescriptive to stimulate discussion in the community. These four challenges are ranked based on their spatial scale: from the challenges requiring a global strategy to those relying on efforts at the national and regional scale.

5.1 Facilitate the creation and increase the comparability of LSH datasets by moving their production to the cloud

We propose that the production of LSH data should be progressively moved to the cloud. Currently, LSH creators download different versions of various data products and process them using different scripts. As an alternative, the relevant datasets should be available in the cloud, together with scripts necessary to process them. Users would upload shapefiles of their catchments and the extraction of hydrometeorological time series and catchment attributes would happen online. This would (i) improve inter-dataset comparability as data products and scripts would be consistent across LSH datasets, (ii) facilitate the production of time series and attributes for new catchments, and (iii) enable the simultaneous update of LSH datasets, for instance when a data product becomes available or covers a longer period. Such a system, accessible and maintained by the community instead of a few individuals, would increase

the perennial nature of LSH datasets, i.e. make them easier to produce and maintain in the mid- to long-term.

We acknowledge that, due to data use restrictions, issues related to ownership, and custodian policies of data providers, some data, such as streamflow data, cannot be uploaded to the cloud. However, metadata and data products derived from these data sets could be cloud-based. Further, there is a growing number of open, global datasets covering a variety of variables relevant for LSH, which can be processed online (see for instance data products involved in Rakovec *et al.* 2016, Addor *et al.* 2017, Beck *et al.* 2017, Nijzink *et al.* 2018). These datasets include terrestrial observations, remotely sensed information, and model-based data, such as reanalyses. Several of these data products, in particular remotely sensed data, are already available on cloud computing portals such as Google Earth Engine¹ and Amazon Web Services.² Alternatively, the data processing may be arranged on a non-commercial data-sharing platform, such as Hydroshare³ or Copernicus.⁴ Initiatives like Pangeo,⁵ which aims to facilitate the use of big data in geosciences, could accelerate the development of cloud-based LSH.

5.2 Coordinate the comparison of data sources to assess their uncertainties and value for hydrological research

The global datasets mentioned above would complement national information that LSH dataset developers have access to. As these global datasets are recent, their reliability and accuracy for different regions of the world is not well characterized yet (Addor *et al.* 2018). Using global datasets alongside better-established regional or national datasets would help to assess their value and limitations for hydrological research and applications. Similarly, the comparison of different data products (e.g. remotely sensed products) using a common cloud-based framework would highlight their differences and uncertainties. Finally, in addition to assessing uncertainties using the spread among products, several products now provide uncertainty estimates for their own data (e.g. Newman *et al.* 2015b, Hengl *et al.* 2017, Cornes *et al.* 2018, Chaney *et al.* 2019), and recent coordinated efforts provide guidance on how to conduct streamflow uncertainty assessments in diverse environments (e.g. Kiang *et al.* 2018). Together, these methods will enable us to better characterize uncertainties in LSH datasets.

5.3 Sustain efforts to characterize human impacts on water systems

The level of detail and diversity of geophysical datasets is increasing rapidly, but the characterization of human impacts is progressing much slower. Although it is difficult to access reliable and consistent water use data, there is an opportunity to use recently released global datasets, such as the Global Reservoir and Dam (GRanD) database (Lehner *et al.* 2011),

¹<https://developers.google.com/earth-engine/datasets/catalog/>.

²<https://registry.opendata.aws>.

³<https://www.hydroshare.org>.

⁴<https://www.copernicus.eu/>.

⁵www.pangeo.io.

global gridded water withdrawals (Huang *et al.* 2018) and land cover datasets to incorporate and classify the various types of water engineering infrastructure and human-induced land changes in LSH datasets. These attributes would help for an improved understanding of human impacts on hydrological catchment functioning over time.

Several authors have stressed the need to develop indices linking water resources and society (Wada *et al.* 2017), including a threshold value to characterize the degree of water scarcity (Falkenmark 1989), the Water Exploitation Index (De Roo *et al.* 2012), the Blue Water Sustainability Indicator (Wada and Bierkens 2014) and a human intervention degree index (Alvarez-Garreton *et al.* 2018). We argue that the inclusion of these and new human intervention indices should be established and standardized, such that meaningful comparisons of human alteration effects can be achieved across catchments globally.

5.4 Increase the accessibility of hydrological data

Large-sample hydrology is a great opportunity to make hydrological research more collaborative. This relies, in particular, on increasing the accessibility of hydrological data. We encourage users to share data that they have produced for basins of LSH datasets, such as new fluxes (e.g. evapotranspiration), new catchment attributes (e.g. land cover change) and also model related data (e.g. parameter sets, model simulations and hydrological projections driven by climate models).

As discussed above, there are currently many restrictions to data sharing because of data ownership issues. This concerns many data types important for the development of LSH, the most crucial being arguably streamflow. The accessibility of streamflow data varies strongly geographically (see Fig. 1). A concerted effort is needed by hydrologists and umbrella organizations (e.g. GRDC, IAHS, WMO) to lobby for the public release of currently inaccessible streamflow datasets (see also discussion in Gupta *et al.* 2014 and in Viglione *et al.*, 2010). This is particularly urgent in regions with little streamflow data readily available, such as south-eastern Asia and central Africa. Technological issues mean that historic data may only be available in hard copy or an outdated format, and resources may be unavailable locally to transcribe or convert it. Thus, financial assistance as part of international collaborations could catalyse data sharing. Furthermore, the WMO has prepared a guide for the rescue of such data, and interested agencies are directed to WMO (2014) for guidance on good practice (see also Brönnimann *et al.* 2018). Clear articulation of local benefits should be outlined. For example, releasing data for inclusion in large-sample datasets ensures that the geographic region is examined by future studies adopting the dataset, yielding operationally significant insights into the regional hydrology at little cost to the nation or agency. This may partially offset the perception that releasing the data means the loss of a strategic asset.

Overall, there is a need to increase the accessibility and comparability of both observed and simulated hydrological time series. The website <http://camels.cr2.cl> (Alvarez-Garreton *et al.* 2018) provides an example of LSH data provision with a high degree of user interaction. We advocate for more

hydrological simulations to be shared, in order to facilitate model comparison, benchmarking and improvement (e.g. Best *et al.* 2015, Newman *et al.* 2017, Kratzert *et al.* 2019). The platform Catch X (<https://ewgis.org/catchx-global/>) provides simulated runoff across 57,646 catchments using global-scale simulations available through the earth2Observe project (Schellekens *et al.* 2017). This recently launched platform also includes other hydrometeorological variables (e.g. evapotranspiration, snowfall, temperature) and land cover information, and could potentially be one of the toolsets to further bridge the gap between LSH and large-scale hydrology.

6 Conclusions

Large-sample hydrology datasets have enabled progress in multiple fields of hydrological sciences, and they are supporting the emergence of novel approaches to better understand water dynamics, relying for instance on machine learning (Section 1). The content and spatial extent of LSH datasets has significantly expanded over the last decade, and the overview provided in Section 2 and Table 2 should help users to select the datasets corresponding best to their needs. Overall, as new mechanisms are implemented to acknowledge datasets in peer-reviewed studies, the recognition of the key role played by datasets in scientific advances is improving.

Yet, we argue here that to sustain the contribution of LSH datasets to hydrological sciences and to widen the scope of LSH studies, it is essential to better coordinate the production of LSH datasets worldwide (Fig. 2). Currently, their use and interpretation is hindered by their lack of comparability, uncertainty estimates and characterization of human impacts, as well as by the still limited access to hydrological data (Section 3). To overcome these limitations, we propose a list of simple actions that can be taken today when producing or updating a LSH dataset (Section 4). Following these guidelines will increase the overall value of LSH datasets for the community. We argue that to truly overcome the challenges LSH is facing, there is also a need for community-wide, longer-term efforts (Section 5). In particular, we propose to move the production of the LSH datasets to the cloud, in order to accelerate their standardization and facilitate their future management.

Following the guidelines and addressing the grand challenges outlined in this paper has the potential to enhance the transparency and reproducibility of hydrological studies, and to lead to better structured, less fragmented LSH datasets. These datasets are necessary to refine our understanding of hydrological processes and model realism, as they enable us to rigorously test hydrological hypotheses and models across a variety of environments (Andréassian *et al.* 2009). LSH datasets have become an essential community resource, they are more complete and diverse than ever, thanks to the contributions of hydrologists and institutions worldwide. Using common LSH datasets, we can increase the comparability of individual studies and, thereby, enhance our ability to learn from their combined results.

Acknowledgements

This study is a contribution to the Large-sample Hydrology working group of the Panta Rhei Research Initiative of the International Association of

Hydrological Sciences (IAHS). We thank Lukas Gudmundsson and the participants of the splinter meeting ‘Large-sample hydrology: facilitating the production and exchange of datasets worldwide’ at the EGU2018 General Assembly for their inputs. Comments from Wolfgang Grabs and an anonymous reviewer are gratefully acknowledged.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

NA is supported by the Swiss National Science Foundation (fellowships P2ZHP2_161963 and P400P2_180791). HXD acknowledges support from the Australia Awards Scholarship (ST000HKE6), the University of Adelaide D R Stranks Fellowship and the University of Michigan Research Fellowship (grant number U064474). CAG is funded by FONDECYT postdoctoral grant no. 3170428 and the Center for Climate and Resilience Research (CR2, CONICYT/FONDAP/15110009). GC is supported by NERC MaRIUS: Managing the Risks, Impacts and Uncertainties of droughts and water Scarcity (grant NE/L010399/1). KF acknowledges the support of the Bureau of Meteorology, Australia (TP705654) and the Australian Research Council (LP170100598). PAM acknowledges support from FONDECYT postdoctoral grant No. 3170079 and CONICYT-PIA Project AFB180004.

ORCID

Nans Addor  <http://orcid.org/0000-0002-6057-3930>

References

- Addor, N., *et al.*, 2017. The CAMELS data set: catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, 21 (10), 5293–5313. doi:10.5194/hess-21-5293-2017
- Addor, N., *et al.*, 2018. A ranking of hydrological signatures based on their predictability in space. *Water Resources Research*, 54, 8792–8812. doi:10.1029/2018WR022606
- Alvarez-Garreton, C., *et al.*, 2018. The CAMELS-CL dataset: catchment attributes and meteorology for large sample studies - Chile dataset. *Hydrology and Earth System Sciences*, 22 (11), 5817–5846. doi:10.5194/hess-22-5817-2018
- Alvarez-Garreton, C., *et al.*, 2019. The impacts of native forests and forest plantations on water supply in Chile. *Forests*, 10 (6), 473. doi:10.3390/f10060473
- Andréassian, V., *et al.*, 2006a. Introduction and synthesis : why should hydrologists work on a large number of basin data sets ? Large sample basin experiments for hydrological model parameterization: results of the Model Parameter Experiment-MOPEX. Vol. 307. CEH Wallingford, UK: IAHS Publ, 1–5.
- Andréassian, V., *et al.*, Eds, 2006b. Large sample basin experiments for hydrological model parameterization: results of the Model Parameter Experiment-MOPEX. Vol. 307. CEH Wallingford, UK: IAHS Publ, 348 + iv.
- Andréassian, V., *et al.*, 2009. HESS opinions “crash tests for a standardized evaluation of hydrological models”. *Hydrology and Earth System Sciences Discussions*, 6, 3669–3685. doi:10.5194/hessd-6-3669-2009
- Andréassian, V., *et al.*, 2014. Seeking genericity in the selection of parameter sets: impact on hydrological model efficiency. *Water Resources Research*, 50, 8356–8366. doi:10.1002/2013WR014761
- Arsenault, R., *et al.*, 2016. CANOPEX: A Canadian hydrometeorological watershed database. *Hydrological Processes*, 30, 2734–2736. doi:10.1002/hyp.10880
- Australian National Data Service (ANDS), Research Data Services (RDS) and the National eResearch Collaboration Tools and Resources Project (NECTAR), 2017. FAIR self assessment tool. Available from: <https://www.ands-nectar-rds.org.au/fair-tool> (accessed 5 Nov 2019).
- Barbarossa, V., *et al.*, 2018. FLO1K, global maps of mean, maximum and minimum annual streamflow at 1 km resolution from 1960 through 2015. *Scientific Data*, 5. doi:10.1038/sdata.2018.78
- Beck, H.E., *et al.*, 2016. Global-scale regionalization of hydrologic model parameters. *Water Resources Research*, 52 (5), 3599–3622. doi:10.1002/2015WR018247
- Beck, H.E., *et al.*, 2017. Global-scale evaluation of 23 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, 21, 6201–6217. doi:10.5194/hess-2017-508
- Beck, H.E., de Roo, A., and van Dijk, A.I.J.M., 2015. Global maps of streamflow characteristics based on observations from several thousand catchments. *Journal of Hydrometeorology*, 16, 1478–1501. doi:10.1175/JHM-D-14-0155.1
- Bennett, B., *et al.*, 2018. An empirical investigation into the effect of antecedent precipitation on flood volume. *Journal of Hydrology*, 567, 435–445. doi:10.1016/j.jhydrol.2018.10.025
- Berghuijs, W.R., *et al.*, 2017. Recent changes in extreme floods across multiple continents. *Environmental Research Letters*, 12 (11), 22–25. doi:10.1088/1748-9326/aa8847
- Best, M.J., *et al.*, 2015. The plumbing of land surface models: benchmarking model performance. *Journal of Hydrometeorology*, 16, 1425–1442. doi:10.1175/JHM-D-14-0158.1
- Bierkens, M.F.P., 2015. Global hydrology: state, trends, and directions. *Water Resources Research*, 51 (7), 4923–4947. doi:10.1002/2015WR017173
- Bierkens, M.F.P., *et al.*, 2015. Hyper-resolution global hydrological modelling: what is next? *Hydrological Processes*, 29, 310–320. doi:10.1002/hyp.10391
- Blenkinsop, S., *et al.*, 2017. Quality control of an hourly rainfall dataset and climatology of extremes for the UK. *International Journal of Climatology*, 37 (2), 722–740. doi:10.1002/joc.4735
- Blöschl, G., *et al.*, 2017. Changing climate shifts timing of European floods. *Science*, 357 (6351), 588–590. doi:10.1126/science.aan2506
- Boldetti, G., *et al.*, 2010. Data-set cleansing practices and hydrological regionalization: is there any valuable information among outliers? *Hydrological Sciences Journal*, 55 (6), 941–951. doi:10.1080/02626667.2010.505171
- Brönnimann, S., *et al.*, 2018. A roadmap to climate data rescue services. *Geoscience Data*, 5 (1), 28–39. doi:10.1002/gdj3.56
- Brown, A.E., *et al.*, 2005. A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. *Journal of Hydrology*, 310 (1–4), 28–61. doi:10.1016/j.jhydrol.2004.12.010
- Burn, D.H. and Whitfield, P.H., 2018. Changes in flood events inferred from centennial length streamflow data records. *Advances in Water Resources*, 121, 333–349. doi:10.1016/j.advwatres.2018.08.017
- Ceola, S., *et al.*, 2015. Virtual laboratories: new opportunities for collaborative water science. *Hydrology and Earth System Sciences*, 19 (4), 2101–2117. doi:10.5194/hess-19-2101-2015
- Chaney, N.W., *et al.*, 2019. POLARIS soil properties: 30-meter probabilistic maps of soil properties over the contiguous United States. *Water Resources Research*, 2018WR022797. doi:10.1029/2018WR022797
- Cloke, H.L. and Hannah, D.M., 2011. Large-scale hydrology: advances in understanding processes, dynamics and models from beyond river basin to global scale. *Hydrological Processes*, 25 (7), 991–995. doi:10.1002/hyp.v25.7
- Cornes, R.C., *et al.*, 2018. An ensemble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research Atmospheres*, 123 (17), 9391–9409. doi:10.1029/2017JD028200
- Coron, L., *et al.*, 2012. Crash testing hydrological models in contrasted climate conditions: an experiment on 216 Australian catchments. *Water Resources Research*, 48, W05552. doi:10.1029/2011WR011721
- Coxon, G., *et al.*, 2013. Diagnostic evaluation of multiple hypotheses of hydrological behaviour in a limits-of-acceptability framework for 24 UK catchments. *Hydrological Processes*, 28 (25), 6135–6150. doi:10.1002/hyp.10096
- Coxon, G., *et al.*, 2015. A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations. *Water Resources Research*, 51 (7), 5531–5546. doi:10.1002/wrcr.v51.7
- De Roo, A., *et al.*, 2012. A multi-criteria optimisation of scenarios for the protection of water resources in Europe. Italy, JRC Scientific and Policy Report, JRC75919, ISSN: 1831-9424, EC-JRC-IES. doi:10.1094/PDIS-11-11-0999-PDN

- Do, H.X., *et al.*, 2018a. The Global Streamflow Indices and Metadata Archive (GSIM) – part 1: the production of a daily streamflow archive and metadata. *Earth System Science Data*, 10, 765–785. doi:10.5194/essd-10-765-2018
- Do, H.X., *et al.*, 2018b. The global streamflow indices and metadata archive - part 1: station catalog and catchment boundary. PANGAEA. doi:10.1594/PANGAEA.887477
- Do, H.X., *et al.* 2019. Historical and future changes in global flood magnitude – evidence from a model-observation investigation. *Hydrology and Earth System Sciences*, in review. doi:10.5194/hess-2019-388
- Do, H.X., Westra, S., and Michael, L., 2017. A global-scale investigation of trends in annual maximum streamflow. *Journal of Hydrology*, 552, 28–43. doi:10.1016/j.jhydrol.2017.06.015
- Duan, Q., *et al.*, 2006. Model Parameter Estimation Experiment (MOPEX): an overview of science strategy and major results from the second and third workshops. *Journal of Hydrology*, 320 (1–2), 3–17. doi:10.1016/j.jhydrol.2005.07.031
- Easterbrook, S.M., 2014. Open code for open science? *Nature Geoscience*, 7 (11), 779–781. doi:10.1038/ngeo2283
- Falkenmark, M., 1989. The massive water scarcity now threatening Africa-why isn't it being addressed? *Ambio*, 18, 112–118.
- Falkenmark, M. and Chapman, T., Eds., 1989. *Comparative hydrology*. Paris: UNESCO.
- Fekete, B.M., Vörösmarty, C.J., and Grabs, W., 2002. High-resolution fields of global runoff combining observed river discharge and simulated water balances. *Global Biogeochemical Cycles*, 16 (3), 15-1–15-10. doi:10.1029/1999GB001254
- Ferguson, I.M. and Maxwell, R.M., 2012. Human impacts on terrestrial hydrology: climate change versus pumping and irrigation. *Environmental Research Letters*, 7 (4), 044022. doi:10.1088/1748-9326/7/4/044022
- Finlayson, B., *et al.*, 1986. World hydrology: a new data base for comparative analysis. *Proceedings of the Hydrology and Water Resources Symposium*, Nat. Conf. Publ. 86/13, Brisbane, Australia. The Institution of Engineers, 288–296.
- Fowler, K., *et al.*, 2016. Simulating runoff under changing climatic conditions: revisiting an apparent deficiency of conceptual rainfall-runoff models. *Water Resources Research*, 52, 1820–1846. doi:10.1002/2015WR018068
- Fowler, K., *et al.*, 2018. Improved rainfall-runoff calibration for drying climate: choice of objective function. *Water Resources Research*, 54, 3392–3408. doi:10.1029/2017WR022466
- Ghiggi, G., *et al.*, 2019. GRUN: an observations-based global gridded runoff dataset from 1902 to 2014. *Earth System Science Data*, 1–32. doi:10.5194/essd-2019-32
- GRDC, 2015. 18–19 June 2014. Koblenz, Germany: Global Runoff Data Centre (GRDC), Report of the Twelfth Meeting of the GRDC Steering Committee, Koblenz, Germany, 23.
- Greene, C.A., and Thirumalai, K., 2019. It's time to shift emphasis away from code sharing. *Eos*, 100. doi:10.1029/2019EO116357
- Gudmundsson, L., *et al.*, 2012. Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe. *Journal of Hydrometeorology*, 13 (2), 604–620. doi:10.1175/JHM-D-11-083.1
- Gudmundsson, L., *et al.*, 2018a. The Global Streamflow Indices and Metadata Archive (GSIM) – part 2: quality control, time-series indices and homogeneity assessment. *Earth System Science Data*, 10, 787–804. doi:10.5194/essd-10-787-2018
- Gudmundsson, L., *et al.*, 2018b. The Global Streamflow Indices and Metadata Archive (GSIM) - part 2: time series indices and homogeneity assessment. PANGAEA. doi:10.1594/PANGAEA.887470
- Gudmundsson, L., *et al.*, 2019. Observed trends in global indicators of mean and extreme streamflow. *Geophysical Research Letters*, 46, 756–766. doi:10.1029/2018GL079725
- Gudmundsson, L. and Seneviratne, S.I., 2016. Observation-based gridded runoff estimates for Europe (E-RUN version 1.1). *Earth System Science Data*, 8 (2), 279–295. doi:10.5194/essd-8-279-2016
- Gupta, H.V., *et al.*, 2014. Large-sample hydrology: a need to balance depth with breadth. *Hydrology and Earth System Sciences*, 18, 463–477. doi:10.5194/hess-18-463-2014
- Hanasaki, N., Kanae, S., and Oki, T., 2006. A reservoir operation scheme for global river routing models. *Journal of Hydrology*, 327, 22–41. doi:10.1016/j.jhydrol.2005.11.011
- Hannah, D.M., *et al.*, 2011. Large-scale river flow archives: importance, current status and future needs. *Hydrological Processes*, 25, 1191–1200. doi:10.1002/hyp.7794
- Harrigan, S., *et al.*, 2018. Benchmarking ensemble streamflow prediction skill in the UK. *Hydrology and Earth System Sciences*, 22, 2023–2039. doi:10.5194/hess-22-2023-2018
- Henck, A.C., *et al.*, 2010. Monsoon control of effective discharge, Yunnan and Tibet. *Geology*, 38, 975–978. doi:10.1130/G31444.1
- Hengl, T., *et al.*, 2017. SoilGrids250m: global gridded soil information based on machine learning. *PLoS One*, 12 (2), e0169748. doi:10.1371/journal.pone.0169748
- Hirpa, F.A., *et al.*, 2018. Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data. *Journal of Hydrology*, 566, 595–606. doi:10.1016/j.jhydrol.2018.09.052
- Horsburgh, J.S., *et al.*, 2009. An integrated system for publishing environmental observations data. *Environmental Modelling & Software*, 24 (8), 879–888. doi:10.1016/j.envsoft.2009.01.002
- Hrachowitz, M., *et al.*, 2013. A decade of Predictions in Ungauged Basins (PUB)—a review. *Hydrological Sciences Journal*, 58 (6), 1198–1255. doi:10.1080/02626667.2013.803183
- Huang, Z., *et al.*, 2018. Reconstruction of global gridded monthly sectoral water withdrawals for 1971–2010 and analysis of their spatiotemporal patterns. *Hydrology and Earth System Sciences*, 22, 2117–2133. doi:10.5194/hess-22-2117-2018
- Hutton, C., *et al.*, 2016. Most computational hydrology is not reproducible, so is it really science?. *Water Resources Research*, 52, 7548–7555. doi:10.1002/2016WR019285
- Jaramillo, F. and Destouni, G., 2015. Local flow regulation and irrigation raise global human water consumption and footprint. *Science*, 350 (6265), 1248–1251. doi:10.1126/science.aad1010
- Kauffeldt, A., *et al.*, 2016. Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environmental Modelling & Software*, 75, 68–76. doi:10.1016/j.envsoft.2015.09.009
- Kiang, J.E., *et al.*, 2018. A comparison of methods for streamflow uncertainty estimation. *Water Resources Research*, 54, 7149–7176. doi:10.1029/2018WR022708
- Knoben, W.J.M., Woods, R.A., and Freer, J.E., 2018. A quantitative hydrological climate classification evaluated with independent streamflow data. *Water Resources Research*, 54. doi:10.1029/2018WR022913
- Kovács, G., 1984. Proposal to construct a coordinating matrix for comparative hydrology. *Hydrological Sciences Journal*, 29 (4), 435–443. doi:10.1080/02626668409490961
- Kratzert, F., *et al.*, 2018. Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*. doi:10.5194/hess-22-6005-2018
- Kratzert, F., *et al.*, 2019. Benchmarking a catchment-aware Long Short-Term Memory Network (LSTM) for large-scale hydrological modeling. *Hydrology and Earth System Sciences Discussions*. doi:10.5194/hess-2019-368
- Kuentz, A., *et al.*, 2017. Understanding hydrologic variability across Europe through catchment classification. *Hydrology and Earth System Sciences*, 21, 2863–2879. doi:10.5194/hess-21-2863-2017
- Lehner, B., 2012. *Derivation of watershed boundaries for GRDC gauging stations based on the hydrosheds drainage network*. Koblenz, Germany: Global Runoff Data Centre in the Federal Institute of Hydrology (BFG), Report 41.
- Lehner, B., *et al.*, 2011. *Global Reservoir and Dam Database, Version 1 (GRanDv1): dams, revision 01*. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). doi:10.7927/H4N877QK
- Lewis, E., *et al.*, 2018. A rule based quality control method for hourly rainfall data and a 1 km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr. *Journal of Hydrology*, 564, 930–943. doi:10.1016/j.jhydrol.2018.07.034
- Liu, X., *et al.*, 2017. Multimodel uncertainty changes in simulated river flows induced by human impact parameterizations. *Environmental Research Letters*, 12, 025009. doi:10.1088/1748-9326/aa5a3a

- Mallakpour, I. and Villarini, G., 2015. The changing nature of flooding across the central United States. *Nature Climate Change*, 5 (3), 250. doi:10.1038/nclimate2516
- Mathevet, T., et al., 2006. A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. *IAHS Red Books Series*, n°307, 211–219.
- McMahon, T., 1982. *Hydrological characteristics of selected rivers of the world*. Paris: Technical documents in hydrology, Unesco.
- McMillan, H., et al., 2016b. Panta Rhei 2013–2015: global perspectives on hydrology, society and change. *Hydrological Sciences Journal*, 61 (7), 1174–1191. doi:10.1080/02626667.2016.1159308
- McMillan, H., Krueger, T., and Freer, J., 2012. Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. *Hydrological Processes*, 26 (26), 4078–4111. doi:10.1002/hyp.v26.26
- McMillan, H.K., Booker, D.J., and Cattoën, C., 2016a. Validation of a national hydrological model. *Journal of Hydrology*, 541, 800–815. doi:10.1016/j.jhydrol.2016.07.043
- Melsen, L., et al., 2018. Mapping (dis)agreement in hydrologic projections. *Hydrology and Earth System Sciences*, 22, 1775–1791. doi:10.5194/hess-22-1775-2018
- Moine, M.P., et al., 2014. Development and exploitation of a controlled vocabulary in support of climate modelling. *Geoscientific Model Development*, 7 (2), 479–493. doi:10.5194/gmd-7-479-2014
- Montanari, A., et al., 2013. Panta Rhei-everything flows: change in hydrology and society – the IAHS scientific decade 2013–2022. *Hydrological Sciences Journal*, 58 (6), 1256–1275. doi:10.1080/02626667.2013.809088
- Newman, A.J., et al., 2015a. Development of a large-sample watershed-scale hydrometeorological dataset for the contiguous USA: dataset characteristics and assessment of regional variability in hydrologic model performance. *Hydrology and Earth System Sciences*, 19, 209–223. doi:10.5194/hess-19-209-2015
- Newman, A.J., et al., 2015b. Gridded ensemble precipitation and temperature estimates for the contiguous United States. *Journal of Hydrometeorology*, 16, 2481–2500. doi:10.1175/JHM-D-15-0026.1
- Newman, A.J., et al., 2017. Benchmarking of a physically based hydrologic model. *Journal of Hydrometeorology*, JHM-D-16-0284.1, 18, 2215–2225. doi:10.1175/JHM-D-16-0284.1
- Nijzink, R.C., et al., 2018. Constraining conceptual hydrological models with multiple information sources. *Water Resources Research*, 54, 8332–8362. doi:10.1029/2017WR021895
- Open Data Charter, 2015. International open data charter. Available from: https://opendatacharter.net/wp-content/uploads/2015/10/open-datacharter-charter_F.pdf (accessed 5 Nov 2019).
- Oudin, L., et al., 2010. Are seemingly physically similar catchments truly hydrologically similar? *Water Resources Research*, 46 (11). doi:10.1029/2009WR008887
- Payan, J.L., et al., 2008. How can man-made water reservoirs be accounted for in a lumped rainfall-runoff model? *Water Resources Research*, 44 (3). doi:10.1029/2007WR005971
- Peel, M., et al., 2001. Identification and explanation of continental differences in the variability of annual runoff. *Journal of Hydrology*, 250, 224–240. doi:10.1016/S0022-1694(01)00438-3
- Peel, M., McMahon, T., and Finlayson, B., 2004. Continental differences in the variability of annual runoff - update and reassessment. *Journal of Hydrology*, 295, 185–197. doi:10.1016/j.jhydrol.2004.03.004
- Perrin, C., et al., 2008. Discrete parameterization of hydrological models: evaluating the use of parameter sets libraries over 900 catchments. *Water Resources Research*, 44 (8). doi:10.1029/2007WR006579
- Rakovec, O., et al., 2016. Multiscale and multivariate evaluation of water fluxes and states over European River Basins. *Journal of Hydrometeorology*, 17 (1), 287–307. doi:10.1175/JHM-D-15-0054.1
- Sawicz, K., et al., 2011. Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrology and Earth System Sciences*, 15, 2895–2911. doi:10.5194/hess-15-2895-2011
- Schaake, J., Cong, S.Z., and Duan, Q.Y., 2006. The US MOPEX data set. *IAHS Publications*, 307, 9–28.
- Schellekens, J., et al., 2017. A global water resources ensemble of hydrological models: the earth2Observe Tier-1 dataset. *Earth System Science Data*, 9, 389–413. doi:10.5194/essd-2016-55
- Schmidt, A.H., et al., 2011. The question of Communist land degradation: new evidence from local erosion and basin-wide sediment yield in SW China and SE Tibet. *Annals of the Association of American Geographers*, 101 (3), 1–19.
- Seibert, J., et al., 2018. Upper and lower benchmarks in hydrological modeling. *Hydrological Processes*, 32, 1120–1125. doi:10.1002/hyp.11476
- Singh, R., Archfield, S.A., and Wagener, T., 2014. Identifying dominant controls on hydrologic parameter transfer from gauged to ungauged catchments - A comparative hydrology approach. *Journal of Hydrology*, 517, 985–996. doi:10.1016/j.jhydrol.2014.06.030
- Sivapalan, M., 2009. The secret to ‘doing better hydrological science’: change the question! *Hydrological Processes*, 23, 1391–1396. doi:10.1002/hyp
- Slater, L.J., et al., 2019. Using R in hydrology: a review of recent developments and future directions. *Hydrology and Earth System Sciences*. doi:10.5194/hess-2019-50
- Slater, L.J. and Villarini, G., 2018. Enhancing the predictability of seasonal streamflow with a statistical-dynamical approach. *Geophysical Research Letters*, 45 (13), 6504–6513. doi:10.1029/2018GL077945
- Stoelzle, M., Stahl, K., and Weiler, M., 2013. Are streamflow recession characteristics really characteristic? *Hydrology and Earth System Sciences*, 17 (2), 817–828. doi:10.5194/hess-17-817-2013
- Thompson, S.E., et al., 2011. Comparative hydrology across AmeriFlux sites: the variable roles of climate, vegetation, and groundwater. *Water Resources Research*, 47 (7), W00J07. doi:10.1029/2010WR009797
- Tijdeman, E., et al., 2018b. Natural and human influences on the link between meteorological and hydrological drought indices for a large set of catchments in the contiguous United States. *Water Resources Research*, 54 (9), 6005–6023. doi:10.1029/2017WR022412
- Tijdeman, E., Bachmair, S., and Stahl, K., 2016. Controls on hydrologic drought duration in near-natural streamflow in Europe and the USA. *Hydrology and Earth System Sciences*, 20 (10), 4043–4059. doi:10.5194/hess-20-4043-2016
- Tijdeman, E., Hannaford, J., and Stahl, K., 2018a. Human influences on streamflow drought characteristics in England and Wales. *Hydrology and Earth System Sciences*, 22 (2), 1051–1064. doi:10.5194/hess-22-1051-2018
- Troch, P.A., et al., 2009. Climate and vegetation water use efficiency at catchment scales. *Hydrological Processes*, 23, 2409–2414. doi:10.1002/hyp.7358
- van Dijk, A.I.J.M., et al., 2013. Global analysis of seasonal streamflow predictability using an ensemble prediction system and observations from 6192 small catchments worldwide. *Water Resources Research*, 49, 2729–2746. doi:10.1002/wrcr.20251
- Veldkamp, T.I.E., et al., 2018. Human impact parameterizations in global hydrological models improve estimates of monthly discharges and hydrological extremes: a multi-model validation study. *Environmental Research Letters*, 13 (5), 055008. doi:10.1088/1748-9326/aab96f
- Vertessy, R.A., 2000. Impacts of plantation forestry on catchment runoff. In: E.K. Sadanandan Nabia and A.G. Brown, eds. *Plantations, farm forestry and water*. Barton, Australia: RIRDC, 9–19, no. 01/20.
- Viglione, A., et al., 2010. Barriers to the exchange of hydrometeorological data in Europe: results from a survey and implications for data policy. *Journal of Hydrology*, 394 (1–2), 63–77. doi:10.1016/j.jhydrol.2010.03.023
- Vitolo, C., et al., 2015. Web technologies for environmental big data. *Environmental Modelling & Software*, 63, 185–198. doi:10.1016/j.envsoft.2014.10.007
- Vörösmarty, C.J., et al., 2000. Global water resources: vulnerability from climate change and population growth. *Science*, 289, 284–288. doi:10.1126/science.289.5477.284
- Wada, Y., et al., 2017. Human–water interface in hydrological modelling: current status and future directions. *Hydrology and Earth System Sciences*, 21, 4169–4193. doi:10.5194/hess-21-4169-2017

- Wada, Y. and Bierkens, M.F., 2014. Sustainability of global water use: past reconstruction and future projections. *Environmental Research Letters*, 9 (10), 104003. doi:[10.1088/1748-9326/9/10/104003](https://doi.org/10.1088/1748-9326/9/10/104003)
- Wehren, B., Schädler, B., and Weingartner, R., 2010. Human interventions. In: U. Bundi, et al., ed. *Alpine waters, the handbook of environmental chemistry*. Vol. 6. Heidelberg: Springer Verlag, 71–92.
- Wilkinson, M.D., et al., 2016. The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. doi:[10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)
- Wood, E.F., et al., 2011. Hyperresolution global land surface modeling: meeting a grand challenge for monitoring earth's terrestrial water. *Water Resources Research*, 47 (5), W05301. doi:[10.1029/2010WR010090](https://doi.org/10.1029/2010WR010090)
- World Meteorological Organization (WMO), 2014. Guidelines for hydrological data rescue. ISBN 978-92-63-11146-3. Switzerland: Geneva 2, WMO-No. 1146. Available online: http://www.wmo.int/pages/prog/hwarp/publications/guidelines_hydrological_DR/wmo_1146_en.pdf
- Zaherpour, J., et al., 2018. Worldwide evaluation of mean and extreme runoff from six global-scale hydrological models that account for human impacts. *Environmental Research Letters*, 13 (6), 065015. doi:[10.1088/1748-9326/aac547](https://doi.org/10.1088/1748-9326/aac547)
- Zhang, L., et al., 2017. Validation of terrestrial water storage variations as simulated by different global numerical models with GRACE satellite observations. *Hydrology and Earth System Sciences*, 21, 821–837. doi:[10.5194/hess-21-821-2017](https://doi.org/10.5194/hess-21-821-2017)
- Zhao, F., et al., 2017. The critical role of the routing scheme in simulating peak river discharge in global hydrological models. *Environmental Research Letters*, 12 (7). doi:[10.1088/1748-9326/aa7250](https://doi.org/10.1088/1748-9326/aa7250).
- Zink, M., et al., 2017. A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty. *Hydrology and Earth System Sciences*, 21, 1769–1790. doi:[10.5194/hess-21-1769-2017](https://doi.org/10.5194/hess-21-1769-2017)